

XGBoost Model for Chronic Kidney Disease Diagnosis

Year: 2020 | Citations: 553 | Authors: Adeola Ogunleye, Qing-Guo Wang

Abstract

Chronic Kidney Disease (CKD) is a menace that is affecting 10 percent of the world population and 15 percent of the South African population. The early and cheap diagnosis of this disease with accuracy and reliability will save 20,000 lives in South Africa per year. Scientists are developing smart solutions with Artificial Intelligence (AI). In this paper, several typical and recent AI algorithms are studied in the context of CKD and the extreme gradient boosting (XGBoost) is chosen as our base model for its high performance. Then, the model is optimized and the optimal full model trained on all the features achieves a testing accuracy, sensitivity, and specificity of 1.000, 1.000, and 1.000, respectively. Note that, to cover the widest range of people, the time and monetary costs of CKD diagnosis have to be minimized with fewest patient tests. Thus, the reduced model using fewer features is desirable while it should still maintain high performance. To this end, the set-theory based rule is presented which combines a few feature selection methods with their collective strengths. The reduced model using about a half of the original full features performs better than the models based on individual feature selection methods and achieves accuracy, sensitivity and specificity, of 1.000, 1.000, and 1.000, respectively.